For AI to succeed, computers need to think more like humans

Imagine if every time you learned something new, you completely forgot how to do a thing you'd already learned.

Finally figured out that taxi-hailing whistle? Now you can't tie your shoes anymore. Learn how to moonwalk; forget how to play the violin. Humans do forget skills, of course, but it usually happens gradually.

Computers forget what they know more dramatically. Learning cannibalizes knowledge. As soon as a new skill is learned, old skills are crowded out. It's a problem computer scientists call "catastrophic forgetting." And it happens because computer brains often rewire themselves—forging new and different connections across neural pathways—every time they learn. This makes it hard for a computer to retain old lessons, but also to learn tasks that require a sequence of steps.

"Researchers will need to solve this problem of catastrophic forgetting for us to get anywhere in terms of producing artificially intelligent computers and robots," said Jeff Clune, an assistant professor of computer science at the University of Wyoming. "Until we do, machines will be mostly one-trick ponies."

So what would it take for a computer brain to retain what it knows, even as it learns new things? That was the question Clune had when he and his colleagues set out to make an artificial brain act more like a human one. Their central idea: See if you can get a computer to organize—and preserve—what it knows within distinct modules of the brain, rather than overwriting what it knows every time it learns something new.

Read full, original article: Teaching a Computer Not to Forget